IJABS 2018: 5:2

Original Article

Validation of Persian Version of PHQ-9 for Diagnosis of Major depressive Episode in Psychiatric Wards in IRAN

Mehdi S Samimi Ardestani [1], Rozita Davari Ashtiani [1*], Zahra Rezaei [2], Sasan Vasegh [3], Shahrokh S Gudarzi [4]

1.Associate professor of psychiatry, behavioral science research center, Imam Hussein medical center, shahid Beheshti university of medical science.
2. private practice psychiatrist .
3. Assistant professor of psychiatry, behavioral science research center, Imam Hussein medical center, shahid Beheshti university of medical science.
4. Consultant Psychiatrist, Melbourne Health, Melbourne VIC.
*(*Corresponding Author: Rzita Davari Ashtiani , Email: rodavari@yahoo.com)*

(Received:10 April 2018 ; Revised: 30 April 2018 ; Accepted:28 May 2018)

Abstract

Introduction: The study was designed to validate a Persian translate of 9-item Patient Health Questionnaire (PHQ-9) for screening depression in clinical setting.
Method: It was a cross sectional study in two psychiatric clinics. 61 Major Depressive Disorder patients (according to Structural Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders), and 61 without this diagnosis entred through convenience sampling. PHQ-9 was given to patients and controls. Intraclass correlation coefficient (ICC and paired T test), Cronbach's alpha coefficient were calculated for test-retest reliability and internal consistency respectively. ROC curves were used to calculate sensitivity and specificity of the test.
Results: According to the ROC curve results, PHQ-9 score equal or greater than 12 offers the greatest Youden Index for the diagnosis of major depression episode (MDE), with a sensitivity and specificity of 80.3% and 78.7%, respectively.
Conclusion: according to our study Persian type of PHQ-9 is a valid and reliable tool for diagnosing major deopressive episodes in clinical settings. It is usefull for detection of MDE severity as well.
Declaration of Interest: None.

Key words: Major depressive disorder, PHQ-9

Introduction

Major depressive disorder is one of the most serious and disabling psychiatric disorders. Major depressive disorder can lead to medical problems, impaired individual functioning, reduced quality of life and mortality (1-3). The prevalence varies among different countries, this discrepancy depends on genetics, environmental parameters, racial factors, cultural differences in perceiving diagnostic parameters, psychometric characteristics of diagnostic tools and sampling protocol (3). The most common meticulous psychiatric disorder in IRAN is major depressive disorder with a 12-month prevalence of 12.7% (4). Major depressive episodes may happen during major depressive disorder (unipolar depression) and also in bipolar disorders both in type I and II .

Despite the high prevalence in community, the rate of detection is very low in all types of depressive disorders so it seems to be reasonable to screen depression especially among populations who have risk factors for developing depression (5-7). According to existing literatures, usage of screening instruments have been proposed as a practical approach in primary care settings for the early detection of depressive disorders (8). In the 1990s, a screening tool named Primary Care

Evaluation of Mental Disorders (PRIME-MD) was developed and validated for diagnosing some types of mental disorders such as, depressive, anxiety, somatoform, alcohol and eating disorders (9). Additionally, a self-directed form of PRIME-MD named, Patient Health Questionnaire (PHQ) was developed for early detection of mental health disorders (10, 11). PHQ-9, GAD-7 and PHQ-15 are three types of PHQ for diagnosis and assessing depression, anxiety disorders and somatic symptom severity respectively (12-14).

PHQ-9 is made up of nine questions according to DSM-IV criteria of major depressive episode (15). It's validity for diagnosis of depression has been proven in many clinical settings including primary (10, 11, 16, 17) and also non-psychiatric (dialysis, cardiology, stroke, diabetes and nephrology) care settings (18-23). This diagnostic tool has been used for diagnosis of depression among different racial and ethnic populations in United States (24). Also, it has been validated in many languages including portuguese (25), german (26, 27), dutch (28, 29), thai (30), malay (31) and konkani (32).

The Beck Depression Inventory is an instrument commonly used for depression screening and is translated to Persian (33).

PHQ-9 is the best instrument for screening of depression in primary care settings because of its shortness and simplicity. It can diagnose depression and measure its severity, so it is suitable for monitoring the treatment process as well (34).

It has been shown that PHQ-9 has a sensitivity of 88% and a specificity of 88% for diagnosing major depression, these two excellent properties besides criteria based ability to diagnosis, make PHQ-9 a very suitable test for clinical and research purposes (15).

The aim of this study was to assess the reliability and validity of the Persian version of the PHQ- 9 screening scale in a clinical setting in Iran.

Methods

Participants

61 patients with the major depressive disorder diagnosis, and 61 without major depressive diagnosis were recrutied from psychiatry wards of Imam Hossein and Taleghani hospitals in Tehran, through convenience sampling.

Inclusion criteria included individuals between 18 and 64 years old, while the exclusion criteria were having psychosis, cognitive impairment or mental retardation.

All patients were selected from patients of the psychiatry ward in their first day of admission while controls were entered from psychiatric outpatient clinic. All the subjects were interviewed according to SCID (Structural Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders) by a psychiatrist to diagnose major depressive episode. After the psychiatric interview, PHQ-9-SR was given to patients with the diagnosis of major depressive episode and the controls. SCID has been validated earlier in Iran (35). Before presenting the participants with the questionnaire, we made sure that all subjects have a writing tool and they were helped to stay mentally relaxed. Then they rendered a written consent and were given a version of PHQ-9-SR along with a demographic questionnaire. All of the subjects were checked to be able to read the questionnaire and had understood that she/he had to read every question carefully and then choose the correct answer. Each subject was asked to choose the closest answer if she/he is not totally sure about it and not to left any question without answer. Before taking back the filled questionnaires, the subjects were asked if they had encountered any vague question; besides, the questionnaires were controlled to check that all questions were answered.

To examine test-retest reliability, 30 of the subjects were asked to answer the self-report scale again after two weeks. In order to assess the patient subjects, psychiatrists interviewed 61 subjects diagnosed with Major Depressive Episode.

In this part of the research, findings of PHQ-9 were compared to the clinical findings based on SCID-I as standard criterion, and psychometric characteristics, sensitivity, specificity, false positive and false negative were calculated.

To respect ethical rules, the aim of research and its confidentiality was explained for the participants and their consent were taken. Besides, treatment was begun for the patients diagnosed with MDE.

Data analysis was done with the help of SPSS-21 software. To calculate test-retest reliability, intraclass correlation coefficient (ICC and paired T test); and to assess internal consistency, Cronbach's alpha coefficient were used. Finally, ROC curves were used to calculate sensitivity and specificity of the test.

Instruments

The main applied tool was PHQ-9 which contains 9 items assessing symptoms mentioned in DSMIV TR (loss of interest, depressed mood, sleep distuebances, reduced energy, change in appetite, worthlessness or guilt feeling, problem in memory and cocentration, psychomotor retardation or aggitation and recurrent thoughts of death or suicide), and one item assessing function. Assessment scale of major depressive episode has been drawn from a 3-page questionnaire called PRIME-MD. Filling this version, takes 3 minutes in 85% of cases and takes 5 minutes in 15% of them (10). PHQ-9 is a self-report questionnaire, but if needed, the items can be read for the participants (9). It is necessary to remind that all items must be answered in one session.

At first, It is translated to the English version of PHQ-9-SR to Persian by a bilingual psychiatrist. Afterwards, it was translated back to English by another bilingual person (not a psychiatrist). This English translation was compared and adopted to the original version by two psychiatrists, and then a poll was conducted among 10 patients referred to psychiatric clinic about fluency and intelligibility of the translation whose result presented no ambiguity in the translated version. Validity of the translated sentences was approved by five faculty member psychiatrists from psychiatric department of Shahid Beheshti University of Medical Science.

Personal information of patiens and controls were collected in the demographic questionaire.

Results

Table 1. Demographic characteristics of the sample are shown in

Table 1. Demographic characteristics of the study sample (N=62)

Characteristics	No.	%
Sex		
Female	62	50.8
Male	60	49.2
Age mean ±SD; Range		34.6 ±10.8; 18-61
Education		
Primary	20	10.9
Guidance	33	17.9
High school	47	25.5
Bachelor and above	23	11.5
Diagnoses		
Bipolar I	28	15.2
Bipolar II	12	6.5
MDD	21	11.4

For a statistical analysis and test-retest measurementreliability,intraclass correlation coefficient was used.

ICC was more than 0.70 for 9 questions and paired T test also showed no significant differences (table 2).

Table 2: pretest-posttest correlation for each question of Persian PHQ-9

Question no	Mean score before	Mean score after	Correlation	ICC*	P value
1	2.8 ± 0.4	1.8 ± 0.9	.558	0.744	.655
2	2.5 ± 0.6	1.5 ± 1	.689	0.816	.414
3	2.1 ± 0.7	1.5 ± 0.9	.729	0.843	.705
4	2.5 ± 0.6	1.6 ± 1	.744	0.744	.414
5	2.2 ± 0.6	1.3 ± 1	.668	0.668	.705
6	2.3 ± 0.7	1.2 ± 1	.741	0.732	.414
7	2.4 ± 0.6	1.3 ± 1	.716	0.715	.257
8	2.2 ± 0.7	1 ± 1	.767	0.751	.414
9	2.4 ± 0.6	1 ± 1	.848	0.838	.025

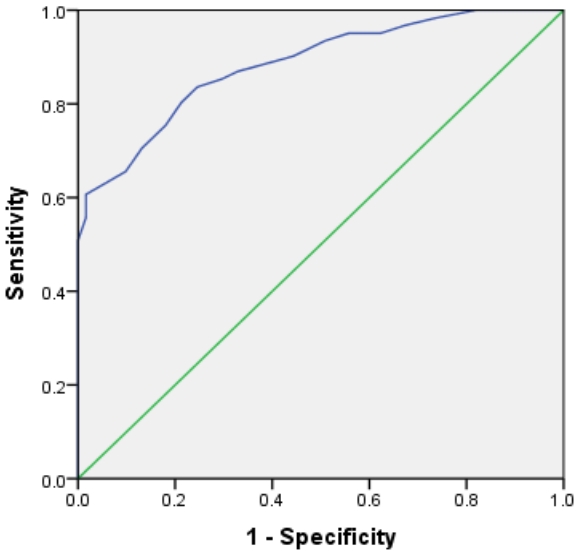
* Intra-class correlation coefficient

Besides, in order to analyze the internal consistency, Cronbach's alpha was calculated for all 9 questions which presented 0.91 to 0.93 for all of them (table 3).

Table no 3: Cronbach's Coefficient Alpha and Item-total correlation for each of questions of Persian PHQ-9.

Question no.	Cronbach's alpha	Item-total correlation
1	.912	.812
2	.918	.713
3	.925	.593
4	.916	.737
5	.918	.714
6	.911	.816
7	.916	.746
8	.916	.737
9	.919	.697

We then draw ROC curve for PHQ-9 results for diagnosing major depressive episode(Figure1).



For calculating sensitivity and specificity with the confidence interval of 95%, after drawing ROC curve for different cut-off points in every 9 question, the highest sensitivity and specificity was found to belong to question 6 with AUC:0.864 and the lowest sensitivity and specificity was found to belong to question 7 with AUC:0.723 .

According to the ROC curve results, PHQ equal or greater than 12 offers the greatest Youden Index for the diagnosis of MDE, with a sensitivity and specificity of 80.3% and 78.7%, respectively.

In table 4 the sensitivity and 1-specificity of various PHQ-9 cut off points for diagnosing major depressive episode that have been illustrated. Often, Youden index is used to determine the best cut off point in the screening questionnaires. As is shown PHQ-9 score equal or greater than 12 offers the greatest Youden Index for the diagnosis of MDE (major depressive episode).

Table 4: Sensitivity and 1-specificity of various PHQ_9 cut off points for Major

Depressive Episode diagnosed by SCID.Cut off score (Than or Equal To)	Sensitivity	1 – Specificity
.5	1.000	.951
1.5	1.000	.902
2.5	1.000	.820
3.5	.984	.738
4.5	.967	.672
5.5	.951	.623
6.5	.951	.557
7.5	.934	.508
8.5	.902	.443
9.5	.869	.328
10.5	.852	.295
11.5	.836	.246
12.5	.803	.213
13.5	.754	.180
14.5	.705	.131
15.5	.656	.098
16.٥	.607	.016
17.5	.557	.016
18.5	.508	٠.000
19.5	.361	٠.000
20.5	.279	٠.000
21.5	.197	٠.000
22.5	.164	٠.000
23.5	.082	٠.000
24.5	.049	٠.000
25.5	.033	٠.000
26.5000	.016	٠.000
28.٠	٠.000	٠.000

A comparison between scores of all 9 questions in patients group and control group showed that there is a significant difference among the scores of two groups. The greatest difference belonged to question 1 which assesses patient's mood regarding depression and decreased pleasure. It means score was 2 and the Inter quartile range (IQR) was obtained as 2-3 in patients groups and 1-2 in control group. The least difference belonged to question 9 which assesses death wish, self-harm and suicidal ideation. It means score was 1 and the IQR was obtained as 1-2 in patients group and 5-1 in controling group. Among subjects of controling group without major depressive episode, through SCID-I interview, 41 subjects were diagnosed with other depressive disorders, and 20 subjects had no kind of depressive disorder. In order to dissociate the cut-off point between the two groups with the confidence interval of 95%, under-curve are for all 9 questions, was less than 0.6 which shows the questionnaire's low ability in discriminating between minor depression and lack of depression in the subjects.

PHQ-9-SR gives two groups of scores (subscales) which are "Major Depressive symptoms" and "impaired function" in patients. A significant statistical correlation between general assessment of functioning (GAF) and disease impact was yeilded with P-value<0.01 and i=0.71. According to the curve, more than 60% of subjects, in GAF higher than 71-80 (impact 1&2) and in 1-10 GAF, had chosen too-much-difficulty in the questionnaire and with the improvement of GAF the difficulty level would be reduced.

Discussion

According to the ROC and AUC cure of this study, PHQ-9 Persian version was suitable for diagnosis major depressive episode.

To our knowledge, this research was the first report of psychometric characteristics of patient health questionnaire 9, self-report version (PHQ-9-SR). One-week test-retest reliability revealed that correlation of all scores in two performance of the questionnaire has been higher than 0.70 ,which correlates with the outcome of the study on English version of the scale (correlation of 0.7-0.8).

A self-report questionnaire is considered reliable only if the Cronbach's alpha is equal or more than 0.70 (36). In our study the Cronbach's alpha was 0.9, which means an appropriate internal consistency similar to surveys accomplished in United States (Cranach's alpha = 0.7-0.89) (24, 37) and Thailand (Cronbach's alpha=0.79) (30).

Like the study that was conducted by Kronke et al., (15) in this study the participants in both groups were assessed by SCID1.

The cutoff point in our study was 12 with the sensitivity of 80.3% and specificity of 78.7%, while in Kronke study the cutoff point was 15 for diagnosis of major depressive episodes the sensitivity and specificity were 87% and 82% respectively (15). A test with sensitivity of 79-97% and specificity of 63-86% is considered acceptable for screening test (38), so Persian version of PHQ-9 is a satisfactory screening test. In other studies for PHQ-9 validation, the obtained cut-off points were for depression as a general syndrome not for Major Depressive Episode (39).

One of the strength of this study was that besides determining the PHQ-9 cutoff for diagnosing major depressive episodes, this scale can be used to determine the severity of major depressive episodes. Also we used for validation SCID-1 as the gold standard, which has a high rate of validity and consistency (35).

To determine the major depressive episode at the cut-off point of 12, satisfactory sensitivity and specificity were obtained (80.3% and 78.7% respectively). However; this questionnaire was unable to screen minor depression from healthy state. Besides, the lowest mean score in each group belonged to question 9 (Thoughts that you would be better off dead or of hurting yourself in some way), this may point out the independency of suicidality from depression. The highest mean score in patients group belonged to question 1 (Little interest or pleasure in doing things). It means that question one (lack of pleasure or anhedonia) is prerequisite for diagnosis of major depressive episode in DSM-IV-TR. The least sensitivity and specificity belonged to

question 7 (Trouble concentrating on things, such as reading the newspaper or watching television) which is not specific for depression and can be observed in other disorders as well. The highest sensitivity and specificity belonged to question 6 (Feeling bad about yourself - or that you are a failure or have let yourself or your family down) which is one of the diagnostic symptoms of depressive disorder and is more specific than other symptoms.

The Persian version of questionnaire was able to measure depression severity. This makes it a suitable tool for comparing different patients suffering from major depressive episodes and evaluating patients longitudinally for example during treatment fallow ups.

In this survey we evaluated the correlation between Persian PHQ-9 and the Global Assessment of Functioning (GAF). The score of PHQ-9 was negatively correlated with GAF in both MDE and non MDE groups ($p<0.001$). It means Persian PHQ-9 can discriminate patients with major depressive episode from ones without, and also this differentiation is functionally meaningful.

According to our study Persian version of PHQ-9 was not able to separate depressive states other than MDE from normal because the controls were someone who may be normal or suffer from mild depression such as minor depressive episode.

Although we yielded a cutoff point with acceptable sensitivity and specificity, this cannot be used for estimating predictive values. The reason is that the participants were selected from clinical settings and not from community.

Like other self-report tools, PHQ-9 is subject to several limitations, for instance gender and cultural differences may affect the answering. Random answering and insufficient motivation would cause prejudice; as well sometimes subjects try to present themselves better. All of these facts are considered as the limitations of the present research.

Acknowledgments

The present research has been conducted with the support of The Center of Behavioral Sciences of Shahid Beheshti University of Medical Science.

We are grateful to all the patients who accompanied us in conducting this research. Also we thanks Dr. Payam Arya M.D. for his kindly supports in gathering articles we needed for this study.

References

1. Ustun TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJ. Global burden of depressive disorders in the year 2000. Br J Psychiatry. 2004 May;184:386-92.
2. Spijker J, Graaf R, Bijl RV, Beekman AT, Ormel J, Nolen WA. Functional disability and depression in the general population. Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). Acta Psychiatr Scand. 2004 Sep;110(3):208-14.
3. Murray CJ, Lopez AD. Evidence-based health policy--lessons from the Global Burden of Disease Study. Science. 1996 Nov 1;274(5288):740-3.
4. Sharifi V, Amin-Esmaeili M, Hajebi A, Motevalian A, Radgoodarzi R, Hefazi M, et al. Twelve-month prevalence and correlates of psychiatric disorders in Iran: the Iranian Mental Health Survey, 2011. *Arch Iran Med*. 2015 Feb;18(2):76-84.
5. Pignone MP, Gaynes BN, Rushton JL, Burchell CM, Orleans CT, Mulrow CD, et al. Screening for depression in adults: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2002 May 21;136(10):765-76.
6. Coyne JC, Schwenk TL, Fechner-Bates S. Nondetection of depression by primary care physicians reconsidered. Gen Hosp Psychiatry. 1995 Jan;17(1):3-12.
7. Simon GE, VonKorff M. Recognition, management, and outcomes of depression in primary care. *Archives of family medicine*. 1995;4(2):99.
8. Mulrow CD, Williams JW, Jr., Gerety MB, Ramirez G, Montiel OM, Kerber C. Case-finding instruments for depression in primary care settings. *Ann Intern Med*. 1995 Jun 15;122(12):913-21.
9. Spitzer RL, Williams JB, Kroenke K, Linzer M, Verloin deGruy F, Hahn SR, et al. Utility of a new procedure for diagnosing mental disorders in primary care: the PRIME-MD 1000 study. *JAMA*. 1994;272(22):1749-56.

10. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA. 1999 Nov 10;282(18):1737-44.
11. Spitzer RL, Williams JB, Kroenke K, Hornyak R, McMurray J. Validity and utility of the PRIME-MD patient health questionnaire in assessment of 3000 obstetric-gynecologic patients: the PRIME-MD Patient Health Questionnaire Obstetrics-Gynecology Study. *Am J Obstet Gynecol*. 2000 Sep;183(3):759-69.
12. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine.* 2006;166(10):1092-7.
13. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Annals of internal medicine. 2007;146(5):317-25.
14. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosomatic medicine. 2002;64(2):258-66.
15. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001 Sep;16(9):606-13.
16. Gilbody S, Richards D, Barkham M. Diagnosing depression in primary care using self-completed instruments: UK validation of PHQ–9 and CORE–OM. Br J Gen Pract. 2007;57(541):650-2.
17. Yeung A, Fung F, Yu S-C, Vorono S, Ly M, Wu S, et al. Validation of the Patient Health Questionnaire-9 for depression screening among Chinese Americans. Comprehensive psychiatry. 2008;49(2):211-7.
18. Fann JR, Bombardier CH, Dikmen S, Esselman P, Warms CA, Pelzer E, et al. Validity of the Patient Health Questionnaire-9 in assessing depression following traumatic brain injury. The Journal of head trauma rehabilitation. 2005;20(6):501-11.
19. Stafford L, Berk M, Jackson HJ. Validity of the Hospital Anxiety and Depression Scale and Patient Health Questionnaire-9 to screen for depression in patients with coronary artery disease. General hospital psychiatry. 2007;29(5):417-24.
20. Thombs BD, Ziegelstein RC, Whooley MA. Optimizing detection of major depression among patients with coronary artery disease using the patient health questionnaire: data from the heart and soul study. Journal of General Internal Medicine. 2008;23(12):2014-7.
21. Williams LS, Brizendine EJ, Plue L, Bakas T, Tu W, Hendrie H, et al. Performance of the PHQ-9 as a screening tool for depression after stroke. stroke. 2005;36(3):635-8.
22. Watnick S, Wang P-L, Demadura T, Ganzini L. Validation of 2 depression screening tools in dialysis patients. American Journal of Kidney Diseases. 2005;46(5):919-24.
23. van Steenbergen-Weijenburg KM, de Vroege L, Ploeger RR, Brals JW, Vloedbeld MG, Veneman TF, et al. Validation of the PHQ-9 as a screening instrument for depression in diabetes patients in specialized outpatient clinics. BMC Health Serv Res. 2010 Aug 12;10:235.
24. Huang FY, Chung H, Kroenke K, Delucchi KL, Spitzer RL. Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. *J Gen Intern Med*. 2006 Jun;21(6):547-52.
25. de Lima Osório F, Vilela Mendes A, Crippa JA, Loureiro SR. Study of the Discriminative Validity of the PHQ-9 and PHQ-2 in a Sample of Brazilian Women in the Context of Primary Health Care. *Perspectives in psychiatric care*. 2009;45(3):216-27.
26. Grafe K, Zipfel S, Herzog W, Lowe B. Screening for psychiatric disorders with the Patient Health Questionnaire (PHQ). Results from the German validation study. Diagnostica. 2004;50(4):171-81.
27. Löwe B, Spitzer RL, Gräfe K, Kroenke K, Quenter A, Zipfel S, et al. Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. *Journal of affective disorders*. 2004;78(2):131-40.
28. Lamers F, Jonkers CC, Bosma H, Penninx BW, Knottnerus JA, van Eijk JTM. Summed score of the Patient Health Questionnaire-9 was a reliable and valid method for depression screening in chronically ill elderly patients. *Journal of clinical epidemiology*. 2008;61(7):679-87.
29. Wittkampf K, van Ravesteijn H, Baas K, van de Hoogen H, Schene A, Bindels P, et al. The accuracy of Patient Health Questionnaire-9 in detecting depression and measuring depression severity in high-risk groups in primary care. *General hospital psychiatry*. 2009;31(5):451-9.

30. Lotrakul M, Sumrithe S, Saipanish R. Reliability and validity of the Thai version of the PHQ-9. *BMC Psychiatry*. 2008 Jun 20;8:46.
31. Azah M, Shah M, Juwita S, Bahri I, Rushidi W, Jamil Y. Validation of the Malay Version Brief Patient Health Questionnaire (PHQ-9) among Adult Attending Family Medicine Clinics. *International Medical Journal*. 2005.
32. Patel V, Araya R, Chowdhary N, King M, Kirkwood B, Nayak S, et al. Detecting common mental disorders in primary care in India: a comparison of five screening questionnaires. *Psychological medicine*. 2008;38(2):221-8.
33. Ghassemzadeh H, Mojtabai R, Karamghadiri N, Ebrahimkhani N. Psychometric properties of a Persian-language version of the Beck Depression Inventory--Second edition: BDI-II-PERSIAN. Depress Anxiety. 2005;21(4):185-92.
34. Nease DE, Jr., Maloin JM. Depression screening: a practical strategy. *J Fam Pract*. 2003 Feb;52(2):118-24.
35. Sharifi V, Assadi SM, Mohammadi MR, Amini H, Kaviani H, Semnani Y, et al. A Persian translation of the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition: psychometric properties. *Compr Psychiatry*. 2009 Jan-Feb;50(1):86-91.
36. Dawson B, Trapp R. Basic and clinical biostatistics 3rd ed. In lange medical books. McGnaw-Hill Medical Publishing Division, New York. 2001.
37. Lee PW, Schulberg HC, Raue PJ, Kroenke K. Concordance between the PHQ-9 and the HSCL-20 in depressed primary care patients. *J Affect Disord*. 2007 Apr;99(1-3):139-45.
38. Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J, et al. The Composite International Diagnostic Interview. An epidemiologic Instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry*. 1988 Dec;45(12):1069-77.
39. Sherina MS, Arroll B, Goodyear-Smith F. Criterion validity of the PHQ-9 (Malay version) in a primary care clinic in Malaysia. *Med J Malaysia*. 2012 Jun;67(3):309-15.